



下载APP



春节加餐01 | 破解还是不破解，这是一个问题

2022-01-31 DS Hunter

《反爬虫兵法演绎20讲》

课程介绍 >



讲述：DS Hunter

时长 10:01 大小 9.18M



你好，我是 DS Hunter。

今天是除夕夜，在这里先祝你新春快乐。

终于远离爬虫和反爬虫尔虞我诈的战场了，有没有轻松一些？今天，我们就先远离这些纷乱的硝烟和危险的斗争，远离爬虫和反爬虫的“知识点”。

我在开篇词的时候就和你说过，希望这门课能够帮助你堂堂正正地，还把工作完成了。而这种虚无缥缈的话，你肯定不信；那些苦口婆心的说教，你肯定也不爱听。所以，咱们来点轻松的，看看在无故事王国里，爬虫和反爬虫的战场是什么样的。



先给你预告一下，春节期间呢，我一共会给你讲三次故事。除夕夜、大年初二的时候每次给你讲两个故事，无故事王国里充满了匪夷所思的事情，不要觉得惊讶。至于看了故事之后给你总结出点什么？故事王国里的事放到现实世界的意义，还是自己理解比较好。爬虫和反爬虫的这些弯弯绕绕，全靠悟。

到了立春那天，也就是大年初四了。我会针对近期的一个爬虫案子给出一点个人的分析。新的一年，咱们看看过去的事，帮着自己在爬虫和反爬虫的这条路上找找方向。当然，这里要注意一点：我是技术人员，不是专业法务，因此所谈的一切，仅限于个人观点，不代表权威法律指南。如果和你们公司法务的说法有冲突，那还是请你以公司法务的说法为准。

那我们现在就开始吧。

故事一

有些事情，不靠技术也可以解决，甚至可能比技术还管用

根据常规故事的开头，这件事应该发生在一个孤独的夜晚，外面交织着青蛙的叫声和潺潺的流水。这是一个偏僻的郊外，创业公司只租得起这样的办公场地。一个爬虫工程师，我们叫他小陶吧，努力睁圆自己小小的眼睛，试图对这个反爬系统做最后的破解。

半个月过去了，小陶对系统进行了大量的破解。刚来的时候，他只会 Python。现在，他对 js (JavaScript) 的作用域、浏览器兼容以及各种 js 黑魔法了如指掌。三天前，公司的前端还来找他请教问题。这一切都来源于在对这个系统的破解中，小陶不断的学习。但是截止到现在，他还是无法拿到 Token 的计算方式。

不过，看起来胜利就在眼前了。他已经顺利地过了四层不同的加解密，代码结构已经变得清晰了，甚至已经解出了一些可读的代码。不出意外的话，天亮之前就能得出结果。

第二天早上，工友发现他睡倒在位置上，嘴角还露出甜甜的微笑，比找到女朋友还开心。

当天下午，他提出了离职申请。

老板大吃一惊：之前毫无征兆啊！这是他们最优秀的工程师，来的时候懵懵懂懂，试用期还没过就已经成长为公司大牛了。但是出于法律要求，也只能批准了他的申请，然后赶紧找人交接。

交接的员工用了半个月，终于也破解到了那个让小陶离开的地方。因为上一个人就是在这里离职的，所以这次大家一起盯着他破解，想看小陶到底看到了什么，居然让他看完就离职了。

交接的员工找到了最后一次破解的关键所在，缓缓敲下最后一行命令，终于看到了真正的代码所在。只见屏幕上赫然写着一行注释：

“当你看到这句话的时候，以你的知识量，应聘一个前端工程师，月薪一般最少是XXXXX元……”

这是一个五位数的月薪，也许在互联网行业比较常见，但是这个创业公司无疑是给不起的。

我们并不清楚后来公司发生了什么，只知道这个交接的人无奈地发了一条知乎来吐槽这件事。

这个事情能告诉我们什么呢？这里并不想过多地引导什么，只能说一个最简单的道理：**有些事情，不靠技术也是可以解决的。甚至，可能比技术还管用。**

此外，员工心理必须时刻关注。很多事情都是心理导致的，而且是长年累月积攒的。等到爆发，其实只需要一个导火索。而一旦爆发，一定是让人措手不及的。

不过，总体来说，还是个值得开心的故事：技术人员能够跳槽、获得更高的薪水，这个结局还是不错的。让我们祝福小陶吧，也祝大家新的一年也能获得高薪。

当然，别忘了，拿高薪的前提，还是要多学习啊。

我们再看下一个小故事。

故事二

你根本不知道自己的哪部分代码在起核心作用

一个公司，爬虫团队一直在试图破解竞对的反爬虫系统。这个反爬虫系统里面有一个很绕的异步循环调用，爬虫团队每次都卡在里面，即使明知道策略发生变化后，这个异步没有变，他们还是会被坑在里面很久，每天都搞得自己要崩溃了。最后，两家握手言和不再竞争了。

这个时候，公司员工和对方交流的时候，忍不住问了一下：你们那个异步循环调用是个什么逻辑？每次我们都被坑在里面。对方愣住了，说：异步循环？我们没有这个东西啊，你说的是什么地方？

然后双方联手调试了下反爬虫代码。

反爬工程师最终感慨地说道：“第一，这是我第一次调试自己的线上代码，真的是太他妈恶心了.....我都要调吐了。第二，你刚刚说的那个异步循环，根本不是什么异步循环，只是一个 Promise 的 polyfill.....因为低版本浏览器不支持 Promise，所以我不得不自己实现一个。至于你问的那个地方为啥感觉跳得很奇怪.....很简单.....我当时没写好，那是个 Bug.....”

这个故事告诉了我们两件事。第一个呢，就是**你根本不知道自己的哪部分代码在起核心作用**。第二个，就在反爬虫代码里，任何 Bug，只要能够让系统正常运行，都将是一个优秀的 Feature。

魔幻吗？不魔幻。

事实上，我们的大部分系统都是这么跑起来的，不仅仅是反爬虫系统，整个系统通常也就是这个样子：



系统非常稳定，所有代码不要随便动！

可想而知，当我们的系统只有依靠一个 Bug 才能正常运行的时候，维护人员内心是多么的崩溃。

而反爬虫写代码则不同，不止自己这面的人要维护，对方的人也要不断地阅读你的代码，这个时候写出一些比较恶心的代码，也许是个好事情呢？

从这个角度出发，我们也可以脑洞大开，难以维护本来是一个贬义词，这里居然成了好事情。那么我们扩展下，其余的研发相关贬义词呢？

性能低、可维护性差，等等。细分下来还有：全局变量、长函数、goto、eval.....

那些你觉得臭名昭著的特性，写一次就被 Leader 骂一次的代码，都可以用上来啊！

当然，我们还是要注意：不要影响自己的阅读。例如 eval，调试的人看到的是 eval，你看到的应该是正常的无 eval 的代码。所以你需要一个转换器。在后面的《10 | 反爬虫概述（中）：无收益的前端如何帮助后端处理爬虫？》里，我还会详细论述具体的做法，你可以期待一下。

小结

那么，今天的故事就到这里了。在别人眼里并不重要的事情，对于当事人来讲，有的时候却是无比重要的。报酬是这样，摇摇欲坠的代码也是这样。

下节加餐，我们会继续以故事为主。敬请期待。

思考题

不用害怕，咱们今天的思考题呢，就没有那么学术了。而我今天，只有一个问题想问你，并且是开放性的。我们都说，反爬虫代码以乱为美，但是这么乱的代码却要嵌入到你的系统里，你的系统也就变乱了，那怎么办呢？

期待你在评论区的分享，我会及时回复。反爬虽然“无定式”，不过春节了，咱们就不一起探索了。我们一起，聊聊天。

分享给需要的人，Ta订阅超级会员，你将得 50 元

Ta单独购买本课程，你将得 20 元

 生成海报并分享

 赞 0  提建议

© 版权归极客邦科技所有，未经许可不得传播售卖。页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

上一篇 03 | 反爬虫的发展以及展望：我们也不是什么正经工程师

下一篇 春节加餐02 | 真实世界里，小心屠龙勇士变恶龙

精选留言 (2)

 写留言



2022-02-02

乱有乱的美啊，反爬代码怎么能达到真正的效果，当然不能只靠乱来取胜啦~

作者回复: 



leslie

2022-01-31

什么地方乱？怎么乱？前几天看军事电视剧提到战争的战法一样，怎么能让对方觉得你乱你又实际不乱。
似乱非乱，形乱神在；不容易啊。

作者回复: 对，就是这种感觉，对方认为乱能迷惑对方，自己不乱方便维护。



共 2 条评论 >

